```
=> fil reg; d que l6

FILE 'REGISTRY' ENTERED AT 15:09:21 ON 14 JUL 2000
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2000 American Chemical Society (ACS)

STRUCTURE FILE UPDATES:   13 JUL 2000  HIGHEST RN 277295-36-0
DICTIONARY FILE UPDATES:  13 JUL 2000  HIGHEST RN 277295-36-0

TSCA INFORMATION NOW CURRENT THROUGH JANUARY 11, 2000

  Please note that search-term pricing does apply when
  conducting SmartSELECT searches.

Structure search limits have been increased.  See HELP SLIMIT
for details.
```

alpha chain — 3 separate query segs combined into one search statement

Position 11 or 13 or 20 forced to change. Other positions may or may not change

```
L6             1 SEA FILE=REGISTRY ABB=ON  APDVQDCPEC[RKH]L[QRHK][ERHK]N[PRHK][F
                 RHK]FS[QRHK]PGAPIL|APDVQDCPEC[TRKH]L[RHK][ERHK]N[PRHK][FRHK]FS[
                 QRHK]PGAPIL|APDVQDCPEC[TRKH]L[QRHK][ERHK]N[PRHK][FRHK]FS[RHK]PG
                 APIL/SQSFP
```

family search - conservative substitution allowed. See sheet attached at end of this part of search.

```
=> d sqd 16

L6   ANSWER 1 OF 1  REGISTRY  COPYRIGHT 2000 ACS
RN   221650-43-7  REGISTRY
FS   PROTEIN SEQUENCE
SQL  92

SEQ      1 APDVQDCPEC TLKKNKFFSK PGAPILQCMG CCFSRAYPTP LRSKKTMLVQ
           ========== ========== ======
        51 KNVTSESTCC VAKSYNRVTV MGGFKVENHT ACHCSTCYYH KS
HITS AT:   1-26
```

RN — use Registry # to match citation to sequence

SQL — sequence length

13 14 16 20

Up scanning sent, agencies?

3 separate query segs combined into one search statement

```
=> d que l7

L7             1 SEA FILE=REGISTRY ABB=ON  APDVQDCPEC[TRKH]L[QRHK][RHK]N[PRHK][F
                 RHK]FS[QRHK]PGAPIL|APDVQDCPEC[TRKH]L[QRHK][ERHK]N[RHK][FRHK]FS[
                 QRHK]PGAPIL|APDVQDCPEC[TRKH]L[QRHK][ERHK]N[PRHK][RHK]FS[QRHK]PG
                 APIL/SQSFP
```

alpha chain — Position 14 or 16 or 17 forced to change. Other positions may or may not change

family search

```
=> d sqd l7

L7   ANSWER 1 OF 1  REGISTRY  COPYRIGHT 2000 ACS
RN   221650-43-7  REGISTRY
FS   PROTEIN SEQUENCE
SQL  92

SEQ      1 APDVQDCPEC TLKKNKFFSK PGAPILQCMG CCFSRAYPTP LRSKKTMLVQ
           ========== ========== ======
        51 KNVTSESTCC VAKSYNRVTV MGGFKVENHT ACHCSTCYYH KS
HITS AT:   1-26

=> d que l10
```

beta chain — substitution at positions 58, 63, 69

```
L10             3 SEA FILE=REGISTRY ABB=ON  VCTYRDF[RHK]YRTV[ERHK]IPGCP[LRHK]HVAP
                  YFSYPVA|VCTYRDF[IRHK]YRTV[RHK]IPGCP[LRHK]HVAPYFSYPVA|VCTYRDF[IR
                  HK]YRTV[ERHK]IPGCP[RHK]HVAPYFSYPVA/SQSFP
```

3 seqs combined in one search statement. — family search

For each, 1 position forced substitution, other 2 may or may not have substitution

```
=> d sqd l10

L10  ANSWER 1 OF 3  REGISTRY  COPYRIGHT 2000 ACS
RN   221650-47-1  REGISTRY
FS   PROTEIN SEQUENCE
SQL  118

SEQ      1 FCIPTEYTMH IERRECAYCL TINTTICAGY CMTRDINGKL FLPKYALSQD
        51 VCTYRDFIYR TVEIPGCPHH VAPYFSYPVA LSCKCGKCNT DYSDCIHEAI
           ========== ========== ==========
       101 KTNYCTKPQK SYLVGFSV
HITS AT:   51-80

=> d l10 2-3 sqd


L10  ANSWER 2 OF 3  REGISTRY  COPYRIGHT 2000 ACS
RN   86179-24-0  REGISTRY
FS   PROTEIN SEQUENCE
SQL  138

SEQ      1 MSAAVLLSVL FALACGQAAS FCIPTEYTMY VDRRECAYCL TINTTICAGY
        51 CMTRDINGKL FLPKYALSQD VCTYRDFIYR TVEIPGCPHH VTPYFSFPVA
                                 ========== ========== ==========
       101 VSCKCGKCNT DNSDCIHEAV RTNYCTKPQS FYLGGFSV
HITS AT:   71-100

L10  ANSWER 3 OF 3  REGISTRY  COPYRIGHT 2000 ACS
RN   86179-23-9  REGISTRY
FS   PROTEIN SEQUENCE
SQL  118

SEQ      1 FCIPTEYTMY VDRRECAYCL TINTTICAGY CMTRDINGKL FLPKYALSQD
        51 VCTYRDFIYR TVEIPGCPHH VTPYFSFPVA VSCKCGKCNT DNSDCIHEAV
           ========== ========== ==========
       101 RTNYCTKPQS FYLGGFSV
HITS AT:   51-80


=> fil capl; s l6 or l7 or l10

FILE 'CAPLUS' ENTERED AT 15:10:18 ON 14 JUL 2000
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2000 AMERICAN CHEMICAL SOCIETY (ACS)
```



```
FILE COVERS 1967 - 14 Jul 2000  VOL 133 ISS 3
FILE LAST UPDATED: 13 Jul 2000  (20000713/ED)
```

This file contains CAS Registry Numbers for easy and accurate substance identification.

This file supports REG1stRY for direct browsing and searching of all substance data from the REGISTRY file. Enter HELP FIRST for more information.

Now you can extend your author, patent assignee, patent information, and title searches back to 1907. The records from 1907-1966 now have this searchable data in CAOLD. You now have electronic access to all of CA: 1907 to 1966 in CAOLD and 1967 to the present in CAPLUS on STN.

```
            1 L6
            1 L7
            4 L10
L14         4 L6 OR L7 OR L10

=> d ibib abs hitrn l14 1-4; fil hom

L14  ANSWER 1 OF 4  CAPLUS  COPYRIGHT 2000 ACS
ACCESSION NUMBER:         1999:223048  CAPLUS
DOCUMENT NUMBER:          130:247459
TITLE:                    Mutants of thyroid stimulating hormone subunits with
                          improved bioactivity and stability
INVENTOR(S):              Weintraub, Bruce D.; Szkudlinski, Mariusz W.
PATENT ASSIGNEE(S):       University of Maryland, Baltimore, USA
SOURCE:                   PCT Int. Appl., 44 pp.
                          CODEN: PIXXD2
DOCUMENT TYPE:            Patent
LANGUAGE:                 English
FAMILY ACC. NUM. COUNT:   1
PATENT INFORMATION:

     PATENT NO.        KIND  DATE          APPLICATION NO.  DATE
     ----------------  ----  --------      ---------------  --------
     WO 9915665         A2   19990401      WO 1998-US19772  19980922
     WO 9915665         A3   19990520
         W:  AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE,
             DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE,
             KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW,
             MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR,
             TT, UA, UG, US, UZ, VN, YU, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
         RW: GH, GM, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES,
             FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI,
             CM, GA, GN, GW, ML, MR, NE, SN, TD, TG
     AU 9894998         A1   19990412      AU 1998-94998    19980922
     EP 1017817         A2   20000712      EP 1998-948422   19980922
         R:  AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT,
             IE, FI
     WO 2000017360      A1   20000330      WO 1999-US5908   19990319
         W:  AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE,
             DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP,
             KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN,
             MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,
             TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD,
             RU, TJ, TM
         RW: GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW, AT, BE, CH, CY, DE, DK,
             ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG,
             CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG
     AU 9931906         A1   20000410      AU 1999-31906    19990319
```

PRIORITY APPLN. INFO.: US 1997-939472 19970922
WO 1998-US19772 19980922
WO 1999-US5908 19990319

AB The present invention is based upon the discovery that mutant .alpha. subunits and mutant .beta. subunits each comprising amino acid substitutions relative to the wild type can be produced and assembled to form a mutant TSH heterodimer or TSH analog that possesses higher bioactivity in vitro and longer half life in vivo. A preferred mutant .alpha. subunit (to be used in conjunction with a modification to increase the serum half-life of the TSH heterodimer having the mutant .alpha. subunit) comprises four mutations: Q13K, E14K, P16K, and Q20K; a preferred mutant .beta. subunit comprises three mutations: I58R, E63R, and L69R. Multiple mutations within a subunit and modifications to increase the half-life of the TSH heterodimer (i.e., .beta.-subunit fusion with the C-terminal extension peptide of human chorionic gonadotropin and/or a .beta. subunit-.alpha. subunit fusion) can act synergistically to achieve bioactivity that is greater than the sum of the increase of the mutations and the long acting modifications. Accordingly, the present invention provides methods for using mutant TSH heterodimers, TSH analogs, fragments, and derivs. thereof for treating or preventing diseases of the thyroid, in particular thyroid cancer. The invention also relates to methods of diagnosis, prognosis and monitoring for thyroid-related functions. Pharmaceutical and diagnostic compns., methods of using mutant TSH heterodimers and TSH analogs with utility for treatment and prevention of metabolic and reproductive diseases are also provided.

IT **221650-43-7P 221650-47-1P**
RL: BAC (Biological activity or effector, except adverse); BPN (Biosynthetic preparation); PRP (Properties); BIOL (Biological study); PREP (Preparation)
(mutants of human TSH subunits with improved bioactivity and stability)

L14 ANSWER 2 OF 4 CAPLUS COPYRIGHT 2000 ACS
ACCESSION NUMBER: 1988:417963 CAPLUS
DOCUMENT NUMBER: 109:17963
TITLE: Organization and nucleotide sequence of the gene encoding the .beta.-subunit of murine thyrotropin
AUTHOR(S): Gordon, David F.; Wood, William M.; Ridgway, E. Chester
CORPORATE SOURCE: Health Sci. Cent., Univ. Colorado, Denver, CO, 80262, USA
SOURCE: DNA (1988), 7(1), 17-26
CODEN: DNAADR; ISSN: 0198-0238
DOCUMENT TYPE: Journal
LANGUAGE: English

AB A murine genomic DNA library was constructed in .lambda. EMBL3 and the nucleotide sequence of the murine TSH .beta.-subunit (TSH.beta.) gene was detd. The cloned gene was derived from a thyrotropic tumor and had no detectable rearrangements when compared to the murine TSH.beta. gene in total genomic DNA. The murine TSH.beta. gene is 5 kb in size and consists of five exons and four introns. The 5' untranslated region of the mRNA is encoded except for a single nucleotide by exons 1, 2, and 3. The protein-coding regions are encoded by exons 4 and 5 while the 3' untranslated region is entirely contained in exon 5. Primer extension anal. using an exon 1-specific primer was used to map the 5' end of the gene. Two transcriptional start sites are present in the murine TSH.beta. gene which appear to be positioned by two TATAAA sequences located 40 bp apart. In all, 99% of transcripts initiate at the downstream site. Transcription from both start sites is affected by thyroidal status in both murine pituitaries and in TtT97 thyrotropic tumors. Finally, sequences homologous with putative-responsive elements and cAMP-responsive elements are present in the 5'-flanking region and may be important in regulating neg. and pos. effects on TSH.beta. gene expression.

Searched by Barb O'Bryen, STIC 308-4291

IT **86179-23-9**, Thyrotropin (mouse .beta.-subunit protein moiety reduced) **86179-24-0**
RL: PRP (Properties)
(amino acid sequence of)

L14 ANSWER 3 OF 4 CAPLUS COPYRIGHT 2000 ACS
ACCESSION NUMBER: 1988:50315 CAPLUS
DOCUMENT NUMBER: 108:50315
TITLE: Expression of the gene for the .beta. subunit of mouse thyrotropin results in multiple mRNAs differing in their 5'-untranslated regions
AUTHOR(S): Wolf, Ofra; Kourides, Ione A.; Gurr, James A.
CORPORATE SOURCE: Lab. Mol. Endocrinol., Meml. Sloan-Kettering Cancer Cent., New York, NY, 10021, USA
SOURCE: J. Biol. Chem. (1987), 262(34), 16596-603
CODEN: JBCHA3; ISSN: 0021-9258
DOCUMENT TYPE: Journal
LANGUAGE: English
AB The gene encoding the .beta. subunit of mouse thyrotropin (TSH.beta.) was isolated from a mouse genomic library, and its nucleotide sequence has been detd. Blot hybridization anal. of restriction enzyme digests of mouse DNA indicates that there is a single mouse TSH.beta. gene. The gene is 4.8 kilobases in length and contains 5 exons, which are 27, 47, 41, 163, and 328 base pairs long. Exons 1, 2, and 3 encode only 5'-untranslated mRNA sequences and are sepd. by introns that are 150 and 380 base pairs long. The protein-coding mRNA sequences are found in exons 4 and 5 and are interrupted by a 460-base-pair intron. The position of this intron, between the codons for amino acids 34 and 35, has been conserved in all the known glycoprotein hormone .beta.-subunit genes. Exons 3 and 4 are sepd. by a large 3.2-kilobase intron. When primer extension anal., using an oligonucleotide primer complementary to exon 4 sequences, was employed to locate the transcription start site, 4 products were obtained. Nucleotide sequencing of these products showed that they were derived from sep. TSH.beta. mRNAs that differed in the lengths of their 5'-untranslated regions. These 5'-untranslated mRNA sequences are derived from different combinations of exons 1, 2, and 3, each spliced to exons 4 and 5. The longest 5'-untranslated sequence, 116 nucleotides long, includes exons 1, 2, and 3 and the first base of exon 4; the shorter 5'-untranslated regions, 75, 69, and 28 nucleotides long, arise by splicing out the second and(or) the third exon sequences. In contrast to the mouse TSH.beta. gene, transcription of the rat TSH.beta. gene from the analogous start site has been reported to give only a single mRNA, with a 5'-untranslated region of 28 nucleotides. Divergence of the mouse and rat TSH.beta. gene sequences at RNA splice sites can account for the absence of exon 2, but not exon 3, sequences in rat TSH.beta. mRNA. Primer extension and RNase protection analyses also showed that the mouse TSB.beta. gene contains a second transcription start site, located 43 base pairs upstream of the first start site, in a position corresponding to that in the rat TSH.beta. gene. Each start site in the mouse gene is flanked by characteristic TATAA box and CAAT box sequences. In the hypothyroid mouse pituitary and in mouse thyrotropic tumors, transcription occurs predominantly from the downstream start site. Elucidation of the structure of the mouse TSH.beta. gene and its 5'-flanking region facilitates the study of the mechanism controlling the prodn. of multiple mouse TSH.beta. mRNAs and the use of the alternative promoters.
IT **86179-23-9 86179-24-0**
RL: PRP (Properties)
(amino acid sequence of)

L14 ANSWER 4 OF 4 CAPLUS COPYRIGHT 2000 ACS
ACCESSION NUMBER: 1983:417486 CAPLUS
DOCUMENT NUMBER: 99:17486
Searched by Barb O'Bryen, STIC 308-4291

TITLE: Cloning of cDNA encoding the pre-.beta. subunit of mouse thyrotropin
AUTHOR(S): Gurr, James A.; Catterall, James F.; Kourides, Ione A.
CORPORATE SOURCE: Lab. Mol. Endocrinol., Mem. Sloan-Kettering Cancer Cent., New York, NY, 10021, USA
SOURCE: Proc. Natl. Acad. Sci. U. S. A. (1983), 80(8), 2122-6
CODEN: PNASA6; ISSN: 0027-8424
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Double-stranded cDNA was synthesized from sucrose gradient-purified poly(A)-contg. mRNA from a mouse thyrotropic tumor, inserted into the PstI site of plasmid pBR322 by using poly(dC).cntdot.poly(dG) homopolymeric extensions, and cloned in Escherichia coli RRI. Plasmids contg. cDNA sequences coding for the .beta. subunit of TSH [9002-71-5] were identified by cell-free translation of hybrid-selected mRNA and immunopptn. with specific antibody to TSH .beta. subunit. Detn. of the nucleotide sequence of 1 cDNA, 595 base pairs in length, allowed deduction of the entire amino acid sequence of the mouse TSH .beta. subunit. The pre-.beta. subunit contains a 20-amino acid N-terminal signal sequence followed by a 118-amino acid mature TSH .beta. subunit. There is 85-90% homol. in amino acid sequence between mouse TSH .beta. subunit and subunits from man, pig, and cow; however, the mouse subunit contains an addnl. 5 and 6 amino acids at its C-terminus compared to the bovine or human and pig subunits, resp. TSH .beta.-subunit mRNA from a mouse thyrotropic tumor was estd. to be 750 nucleotides in length by hybridization with labeled TSH .beta.-subunit cDNA.

IT **86179-23-9 86179-24-0**
RL: PRP (Properties)
(amino acid sequence of)

FILE 'HOME' ENTERED AT 15:10:32 ON 14 JUL 2000

Sequence Family Search of Proteins (/sqsfp)

In the sequence family search, each amino acid in the query has to match either the exact amino acid or a family member equivalent, as shown in the Family Equivalence Table below. The Family Equivalence Table is applied only to each common amino acid in the sequence. Specific uncommon amino acids may be included in the sequence; however, family equivalents only exist for the common amino acids. An amino acid family is based on a conservative substitution of amino acids sharing a similar chemical property. Each common amino acid in the query is converted to its family class members in a search. A match occurs on a query sequence if each amino acid is exactly matched or any of its family members are encountered. For example, the Hydrophobic-Aromatic family consists of the common amino acids F, W, and Y. If the amino acid F is specified within a sequence exact family search, it will match on amino acids F, W, or Y.

FAMILY EQUIVALENCE TABLE

Family Class Name	**Family Class Members**
Neutral-Weakly Hydrophobic	Ala (A), Gly (G), Pro (P), Ser (S), Thr (T)
Hydrophilic-Acid Amine	Asn (N), Asp (D), Gln (Q), Glu (E)
Hydrophilic-Basic	Arg (R), His (H), Lys (K)
Hydrophobic	Ile (I), Met (M), Leu (L), Val (V)
Hydrophobic-Aromatic	Phe (F), Trp (W), Tyr (Y)
Crosslinking	Cys (C)